EXHIBIT A

Fee Schedule

TABLE-1: LIQUIDATED DAMAGES TABLE FOR TIER-1 MEASURES

PER AFFECTED ITEM						
	Month 1	Month 2	Month3	Month4	Month 5	Month 6
Pre-Ordering	$20	$30	$40	$50	$60	$70
Ordering	$40	$50	$60	$70	$80	$90
Provisioning	$100	$125	$175	$250	$325	$500
Provisioning UNE (Coordinated Customer Conversions)	$400	$450	$500	$550	$650	$800
Maintenance and Repair	$100	$125	$175	$250	$325	$500
Maintenance and Repair UNE	$400	$450	$500	$550	$650	$800
LNP	$150	$250	$500	$600	$700	$800
Billing	$1.00	$1.00	$1.00	$1.00	$1.00	$1.00
IC Trunks	$100	$125	$175	$250	$325	$500
Collocation	$5,000	$5,000	$5,000	$5,000	$5,000	$5,000

TABLE-2: REMEDY PAYMENTS FOR TIER-2 MEASURES

	Per Affected Item
OSS Pre-Ordering	$20
Ordering	$60
Provisioning	$300
Provisioning-UNE (Coordinated Customer Conversions)	$875
Maintenance and Repair	$300
Maintenance and Repair-UNE	$875
Billing	$1.00
LNP	$500
IC Trunks	$500
Collocation	$15,000
Change Management	$1,000

EXHIBIT B

SEEM Sub-Metrics

SEEM TIER-1 SUB-METRICS

1. Reject Interval - Fully Mechanized, Partially Mechanized, Non-Mechanized
2. Firm Order Confirmation Timeliness - Fully Mechanized, Partially Mechanized, Non-Mechanized
3. Percent Missed Installation Appointments - Resale POTS
4. Percent Missed Installation Appointments - Resale Design
5. Percent Missed Installation Appointments - UNE Loop
6. Percent Missed Installation Appointments - UNE Loop & Port Combo
7. Percent Missed Installation Appointments - UNE xDSL (ADSL, HDSL, UCL)
8. Percent Missed Installation Appointments - UNE Line Sharing
9. Percent Missed Installation Appointments - Interconnection Trunks
10. Average Order Completion Interval - Resale POTS
11. Average Order Completion Interval - Resale Design
12. Average Order Completion Interval - UNE Loop
13. Average Order Completion Interval - UNE Loop & Port Combo
14. Average Order Completion Interval - UNE xDSL (ADSL, HDSL, UCL)
15. Average Order Completion Interval - Line Sharing
16. Average Order Completion Interval - Interconnection Trunks
17. Average Completion Notice Interval
18. Coordinated Customer Conversions Interval
19. Coordinated Customer Conversion - Hot Cut Timeliness Percent within Interval and Average Interval
20. Percent Provisioning Troubles within 30 Days - Resale POTS
21. Percent Provisioning Troubles within 30 Days - Resale Design
22. Percent Provisioning Troubles within 30 Days - UNE Loop
23. Percent Provisioning Troubles within 30 Days - UNE Loop & Port Combo
24. Percent Provisioning Troubles within 30 Days - UNE xDSL (ADSL, HDSL, UCL)
25. Percent Provisioning Troubles within 30 Days - UNE Line Sharing
26. Percent Provisioning Troubles within 30 Days - Interconnection Trunks
27. LNP - Percent Missed Installation Appointments
28. LNP – Average Disconnect Timeliness Interval & Disconnect Timeliness Interval Distribution
29. Missed Repair Appointments - Resale POTS
30. Missed Repair Appointments - Resale Design
31. Missed Repair Appointments - UNE Loop + Port Combo
32. Missed Repair Appointments - UNE Loops
33. Missed Repair Appointments - UNE xDSL
34. Missed Repair Appointments - UNE Line Sharing
35. Missed Repair Appointments - Interconnection Trunks
36. Customer Trouble Report Rate - Resale POTS
37. Customer Trouble Report Rate - Resale Design
38. Customer Trouble Report Rate - UNE Loop + Port Combinations
39. Customer Trouble Report Rate - UNE Loops

SEEM TIER-1 SUB-METRICS
CONTINUED

40. Customer Trouble Report Rate - UNE xDSL
41. Customer Trouble Report Rate - UNE Line Sharing
42. Customer Trouble Report Rate - Local Interconnection Trunks
43. Maintenance Average Duration - Resale POTS
44. Maintenance Average Duration - Resale Design
45. Maintenance Average Duration - UNE Loop + Port Combinations
46. Maintenance Average Duration - UNE Loops
47. Maintenance Average Duration - UNE xDSL
48. Maintenance Average Duration - UNE Line Sharing
49. Maintenance Average Duration - Local Interconnection Trunks
50. Percent Repeat Troubles within 30 Days - Resale POTS
51. Percent Repeat Troubles within 30 Days - Resale Design
52. Percent Repeat Troubles within 30 Days - UNE Loop + Port Combinations
53. Percent Repeat Troubles within 30 Days - UNE Loops
54. Percent Repeat Troubles within 30 Days - UNE xDSL
55. Percent Repeat Troubles within 30 Days - UNE Line Sharing
56. Percent Repeat Troubles within 30 Days - Local Interconnection Trunks
57. Trunk Group Performance – CLEC Specific
58. Collocation Percent of Due Dates Missed

SEEM TIER-2 SUB-METRICS

1. Average Response Time and Response Interval
2. Interface Availability
3. Loop Makeup Inquiry - Manual
4. Loop Makeup Inquiry - Electronic
5. Flow Through Service Request (Summary)
6. Reject Interval
7. Firm Order Confirmation Timeliness
8. Percent Missed Installation Appointments - Resale POTS
9. Percent Missed Installation Appointments - Resale Design
10. Percent Missed Installation Appointments - UNE Loop
11. Percent Missed Installation Appointments - UNE Loop & Port Combo
12. Percent Missed Installation Appointments - UNE xDSL (ADSL, HDSL, UCL)
13. Percent Missed Installation Appointments - UNE Line Sharing
14. Percent Missed Installation Appointments - Interconnection Trunks
15. Average Order Completion Interval - Resale POTS
16. Average Order Completion Interval - Resale Design
17. Average Order Completion - UNE Loop
18. Average Order Completion - UNE Loop & Port Combo
19. Average Order Completion - UNE xDSL (ADSL, HDSL, UCL)
20. Average Order Completion - Line Sharing
21. Average Order Completion - Interconnection Trunks
22. Average Completion Notice Interval
23. Coordinated Customer Conversions Interval
24. Coordinated Customer Conversion - Hot Cut Timeliness Percent within Interval and Average Interval
25. Percent Provisioning Troubles within 30 Days - Resale POTS
26. Percent Provisioning Troubles within 30 Days - Resale Design
27. Percent Provisioning Troubles within 30 Days - UNE Loop
28. Percent Provisioning Troubles within 30 Days - UNE Loop & Port Combo
29. Percent Provisioning Troubles within 30 Days - UNE xDSL (ADSL, HDSL, UCL)
30. Percent Provisioning Troubles within 30 Days - UNE Line Sharing
31. Percent Provisioning Troubles within 30 Days - Interconnection Trunks
32. LNP Percent Missed Installation Appointments
33. LNP – Average Disconnect Timeliness Interval & Disconnect Timeliness Interval Distribution
34. Missed Repair Appointments - Resale POTS
35. Missed Repair Appointments - Resale Design
36. Missed Repair Appointments - UNE Loop + Port Combo
37. Missed Repair Appointments - UNE Loops
38. Missed Repair Appointments - UNE xDSL
39. Missed Repair Appointments - UNE Line Sharing

SEEM TIER-2 SUB-METRICS
CONTINUED

40. Missed Repair Appointments - Interconnection Trunks
41. Customer Trouble Report Rate - Resale POTS
42. Customer Trouble Report Rate - Resale Design
43. Customer Trouble Report Rate - UNE Loop + Port Combinations
44. Customer Trouble Report Rate - UNE Loops
45. Customer Trouble Report Rate - UNE xDSL
46. Customer Trouble Report Rate - UNE Line Sharing
47. Customer Trouble Report Rate - Local Interconnection Trunks
48. Maintenance Average Duration - Resale POTS
49. Maintenance Average Duration - Resale Design
50. Maintenance Average Duration - UNE Loop + Port Combinations
51. Maintenance Average Duration - UNE Loops
52. Maintenance Average Duration - UNE xDSL
53. Maintenance Average Duration - UNE Line Sharing
54. Maintenance Average Duration - Local Interconnection Trunks
55. Percent Repeat Troubles within 30 Days - Resale POTS
56. Percent Repeat Troubles within 30 Days - Resale Design
57. Percent Repeat Troubles within 30 Days - UNE Loop + Port Combinations
58. Percent Repeat Troubles within 30 Days - UNE Loops
59. Percent Repeat Troubles within 30 Days - UNE xDSL
60. Percent Repeat Troubles within 30 Days - UNE Line Sharing
61. Percent Repeat Troubles within 30 Days - Local Interconnection Trunks
62. Invoice Accuracy
63. Mean time to Deliver Invoices
64. Usage Data Delivery Accuracy
65. Usage Data Delivery Timeliness
66. Percent Trunk Performance – Aggregate
67. Percent of Due Dates Missed
68. Change Management Notices Sent on Time

<u>SEEM TIER-3 SUB-METRICS</u>

1. Percent Missed Installation Appointments - Resale POTS
2. Percent Missed Installation Appointments - Resale Design
3. Percent Missed Installation Appointments - UNE Loop
4. Percent Missed Installation Appointments - UNE Loop & Port Combo
5. Percent Missed Installation Appointments - UNE xDSL (ADSL, HDSL, UCL)
6. Percent Missed Installation Appointments - UNE Line Sharing
7. Percent Missed Installation Appointments - Interconnection Trunks
8. Missed Repair Appointments - Resale POTS
9. Missed Repair Appointments - Resale Design
10. Missed Repair Appointments - UNE Loop + Port Combo
11. Missed Repair Appointments - UNE Loops
12. Missed Repair Appointments - UNE xDSL
13. Missed Repair Appointments - UNE Line Sharing
14. Missed Repair Appointments - Interconnection Trunks
15. Invoice Accuracy
16. Mean Time To Deliver Invoices
17. Trunk Group Performance – Aggregate
18. Collocation Percent of Due Dates Missed

EXHIBIT C

Statistical Methodology

Statistical Methods for BellSouth Performance Measure Analysis

I. Necessary Properties for a Test Methodology

The statistical process for testing if competing local exchange carriers (CLECs) customers are being treat equally with BellSouth (BST) customers involves more than just a mathematical formula. Three key elements need to be considered before an appropriate decision process can be developed. These are:

- the type of data,
- the type of comparison, and
- the type of performance measure.

Once these elements are determined a test methodology should be developed that complies with the following properties.

- Like-to-Like Comparisons. When possible, data should be compared at appropriate levels, e.g. wire center, time of month, dispatched, residential, new orders. The testing process should:
  - Identify variables that may affect the performance measure.
  - Record these important confounding covariates.
  - Adjust for the observed covariates in order to remove potential biases and to make the CLEC and the ILEC units as comparable as possible.
- Aggregate Level Test Statistic. Each performance measure of interest should be summarized by one overall test statistic giving the decision maker a rule that determines whether a statistically significant difference exists. The test statistic should have the following properties.
  - The method should provide a single overall index, on a standard scale.
  - If entries in comparison cells are exactly proportional over a covariate, the aggregated index should be very nearly the same as if comparisons on the covariate had not been done.
  - The contribution of each comparison cell should depend on the number of observations in the cell.
  - Cancellation between comparison cells should be limited.
  - The index should be a continuous function of the observations.
- Production Mode Process. The decision system must be developed so that it does not require intermediate manual intervention, i.e. the process must be a "black box."
  - Calculations are well defined for possible eventualities.

- The decision process is an algorithm that needs no manual intervention.
- Results should be arrived at in a timely manner.
- The system must recognize that resources are needed for other performance measure-related processes that also must be run in a timely manner.
- The system should be auditable, and adjustable over time.

- Balancing. The testing methodology should balance Type I and Type II Error probabilities.
  - P(Type I Error) = P(Type II Error) for well defined null and alternative hypotheses.
  - The formula for a test's balancing critical value should be simple enough to calculate using standard mathematical functions, i.e. one should avoid methods that require computationally intensive techniques.
  - Little to no information beyond the null hypothesis, the alternative hypothesis, and the number of observations should be required for calculating the balancing critical value.

- Trimming. Trimming of extreme observations from BellSouth and CLEC distributions is needed in order to ensure that a fair comparison is made between performance measures. Three conditions are needed to accomplish this goal. These are:
  - Trimming should be based on a general rule that can be used in a production setting.
  - Trimmed observations should not simply be discarded; they need to be examined and possibly used in the final decision making process.
  - Trimming should only be used on performance measures that are sensitive to "outliers."

Measurement Types

The performance measures that will undergo testing are of four types:

1) means
2) proportions,
3) rates, and
4) ratio

While all four have similar characteristics, proportions and rates are derived from count data while means and ratios are derived from interval measurements.

II. Testing Methodology – The Truncated Z

Many covariates are chosen in order to provide deep comparison levels. In each comparison cell, a Z statistic is calculated. The form of the Z statistic may vary depending on the performance measure, but it should be distributed approximately as a standard normal, with mean zero and variance equal to one. Assuming that the test statistic is derived so that it is negative when the performance for the CLEC is worse than for the ILEC, a positive truncation is done – i.e. if the result is negative it is left alone, if the result is positive it is changed to zero. A weighted average of the truncated statistics is calculated where a cell-weight depends on the volume of BST and CLEC orders in the cell. The weighted average is re-centered by the theoretical mean of a truncated distribution, and this is divided by the standard error of the weighted average. The standard error is computed assuming a fixed effects model.

Proportion Measures

For performance measures that are calculated as a proportion, in each adjustment cell, the truncated Z and the moments for the truncated Z can be calculated in a direct manner. In adjustment cells where proportions are not close to zero or one, and where the sample sizes are reasonably large, a normal approximation can be used. In this case, the moments for the truncated Z come directly from properties of the standard normal distribution. If the normal approximation is not appropriate, then the Z statistic is calculated from the hypergeometric distribution. In this case, the moments of the truncated Z are calculated exactly using the hypergeometric probabilities.

Rate Measures

The truncated Z methodology for rate measures has the same general structure for calculating the Z in each cell as proportion measures. For a rate measure, there are a fixed number of circuits or units for the CLEC, n_{2j} and a fixed number of units for BST, n_{1j}. Suppose that the performance measure is a "trouble rate." The modeling assumption is that the occurrence of a trouble is independent between units and the number of troubles in n circuits follows a Poisson distribution with mean λ n where λ is the probability of a trouble in 1 circuit and n is the number of circuits.

In an adjustment cell, if the number of CLEC troubles is greater than 15 and the number of BST troubles is greater than 15, then the Z test is calculated using the normal approximation to the Poisson. In this case, the moments of the truncated Z come directly from properties of the standard normal distribution. Otherwise, if there are very few troubles, the number of CLEC troubles can be modeled using a binomial distribution with n equal to the total number of troubles (CLEC plus BST troubles.) In this case, the moments for the truncated Z are calculated explicitly using the binomial distribution.

Mean Measures

For mean measures, an adjusted t statistic is calculated for each like-to-like cell, which has at least 7 BST and 7 CLEC transactions. A permutation test is used when one or both of the BST and CLEC sample sizes is less than 6. Both the adjusted t statistic and the permutation calculation are described in the technical description section.

Ratio Measures

Rules will be given for computing a cell test statistic for a ratio measure, however, the current plan for measures in this category, namely billing accuracy, does not call for the use of a Z parity statistic.

III. Technical Description

We start by assuming that any necessary trimming[1] of the data is complete, and that the data are disaggregated so that comparisons are made within appropriate classes or adjustment cells that define "like" observations.

Notation and Exact Testing Distributions

Below, we have detailed the basic notation for the construction of the truncated z statistic. In what follows the word "cell" should be taken to mean a like-to-like comparison cell that has both one (or more) ILEC observation and one (or more) CLEC observation.

L = the total number of occupied cells

j = $1,\ldots,L$; an index for the cells

n_{1j} = the number of ILEC transactions in cell j

n_{2j} = the number of CLEC transactions in cell j

n_j = the total number transactions in cell j; $n_{1j}+ n_{2j}$

X_{1jk} = individual ILEC transactions in cell j; $k = 1,\ldots, n_{1j}$

X_{2jk} = individual CLEC transactions in cell j; $k = 1,\ldots, n_{2j}$

Y_{jk} = individual transaction (both ILEC and CLEC) in cell j

[1] When it is determined that a measure should be trimmed, a trimming rule that is easy to implement in a production setting is:

> **Trim the ILEC observations to the largest CLEC value from all CLEC observations in the month under consideration.**

That is, no CLEC values are removed; all ILEC observations greater than the largest CLEC observation are trimmed.

$$= \begin{cases} X_{1jk} & k = 1,\ldots,n_{1j} \\ X_{2jk} & k = n_{1j}+1,\ldots,n_j \end{cases}$$

$\Phi^{-1}(\cdot)$ = the inverse of the cumulative standard normal distribution function

For Mean Performance Measures the following additional notation is needed.

$\overline{X}_{1j}$ = The ILEC sample mean of cell j

$\overline{X}_{2j}$ = The CLEC sample mean of cell j

s^2_{1j} = The ILEC sample variance in cell j

s^2_{2j} = The CLEC sample variance in cell j

$\{y_{jk}\}$ = a random sample of size n_{2j} from the set of $Y_{j1},\ldots,Y_{jn_j}$; $k = 1,\ldots,n_{2j}$

M_j = The total number of distinct pairs of samples of size n_{1j} and n_{2j};

$$= \binom{n_j}{n_{1j}}$$

The exact parity test is the permutation test based on the "modified Z" statistic. For large samples, we can avoid permutation calculations since this statistic will be normal (or Student's t) to a good approximation. For small samples, where we cannot avoid permutation calculations, we have found that the difference between "modified Z" and the textbook "pooled Z" is negligible. We therefore propose to use the permutation test based on pooled Z for small samples. This decision speeds up the permutation computations considerably, because for each permutation we need only compute the sum of the CLEC sample values, and not the pooled statistic itself.

A permutation probability mass function distribution for cell j, based on the "pooled Z" can be written as

$$PM(t) = P(\sum_k y_{jk} = t) = \frac{\textit{the number of samples that sum to } t}{M_j},$$

and the corresponding cumulative permutation distribution is

$$CPM(t) = P(\sum_k y_{jk} \le t) = \frac{\textit{the number of samples with sum } \le t}{M_j}.$$

For Proportion Performance Measures the following notation is defined

a_{1j} = The number of ILEC cases possessing an attribute of interest in cell j

a_{2j} = The number of CLEC cases possessing an attribute of interest in cell j

a_j = The number of cases possessing an attribute of interest in cell j; $a_{1j} + a_{2j}$

The exact distribution for a parity test is the hypergeometric distribution. The hypergeometric probability mass function distribution for cell j is

$$HG(h) = P(H = h) = \begin{cases} \dfrac{\binom{n_{1j}}{h}\binom{n_{2j}}{a_j - h}}{\binom{n_j}{a_j}}, & \max(0, a_j - n_{2j}) \le h \le \min(a_j, n_{1j}) \\ 0 & \text{otherwise} \end{cases},$$

and the cumulative hypergeometric distribution is

$$CHG(x) = P(H \le x) = \begin{cases} 0 & x < \max(0, a_j - n_{2j}) \\ \displaystyle\sum_{h=\max(0, a_j - n_{1j})}^{x} HG(h), & \max(0, a_j - n_{2j}) \le x \le \min(a_j, n_{1j}) \\ 1 & x > \min(a_j, n_{1j}) \end{cases}.$$

For Rate Measures, the notation needed is defined as

b_{1j} = The number of ILEC base elements in cell j

b_{2j} = The number of CLEC base elements in cell j

b_j = The total number of base elements in cell j; $b_{1j} + b_{2j}$

$\hat{r}_{1j}$ = The ILEC sample rate of cell j; n_{1j}/b_{1j}

$\hat{r}_{2j}$ = The CLEC sample rate of cell j; n_{2j}/b_{2j}

q_j = The relative proportion of ILEC elements for cell j; b_{1j}/b_j

The exact distribution for a parity test is the binomial distribution. The binomial probability mass function distribution for cell j is

$$BN(k) = P(B = k) = \begin{cases} \binom{n_j}{k} q_j^k (1 - q_j)^{n_j - k}, & 0 \le k \le n_j \\ 0 & \text{otherwise} \end{cases},$$

and the cumulative binomial distribution is

$$\mathrm{CBN}(x) = P(B \le x) = \begin{cases} 0 & x < 0 \\ \sum_{k=0}^{x} \mathrm{BN}(k), & 0 \le x \le n_j \\ 1 & x > n_j \end{cases}.$$

For Ratio Performance Measures the following additional notation is needed.

U_{1jk} = additional quantity of interest of an individual ILEC transaction in cell j; k = $1,\ldots, n_{1j}$

U_{2jk} = additional quantity of interest of an individual CLEC transaction in cell j; k = $1,\ldots, n_{2j}$

$\hat{R}_{ij}$ = the ILEC (I = 1) or CLEC (i = 2) ratio of the total additional quantity of interest to the base transaction total in cell j, i.e., $\sum_k U_{ijk} \Big/ \sum_k X_{ijk}$

Calculating the Truncated Z

The general methodology for calculating an aggregate level test statistic is outlined below.

1. **Calculate cell weights**, W_j. A weight based on the number of transactions is used so that a cell, which has a larger number of transactions, has a larger weight. The actual weight formulae will depend on the type of measure.

Mean or Ratio Measure

$$W_j = \sqrt{\frac{n_{1j}n_{2j}}{n_j}}$$

Proportion Measure

$$W_j = \sqrt{\frac{n_{2j}n_{1j}}{n_j} \cdot \frac{a_j}{n_j} \cdot \left(1 - \frac{a_j}{n_j}\right)}$$

Rate Measure

$$W_j = \sqrt{\frac{b_{1j}b_{2j}}{b_j} \cdot \frac{n_j}{b_j}}$$

2. **In each cell, calculate a Z value,** Z_j. A Z statistic with mean 0 and variance 1 is needed for each cell.

- If $W_j = 0$, set $Z_j = 0$.
- Otherwise, the actual Z statistic calculation depends on the type of performance measure.

Mean Measure

$$Z_j = \Phi^{-1}(\alpha)$$

where α is determine by the following algorithm.

If $\min(n_{1j}, n_{2j}) > 6$, then determine α as

$$\alpha = P(t_{n_{1j}-1} \le T_j),$$

that is, α is the probability that a t random variable with n_{1j} - 1 degrees of freedom, is less than

$$T_j = \begin{cases} t_j + \dfrac{g}{6}\left(\dfrac{n_{1j}+2n_{2j}}{\sqrt{n_{1j}\,n_{2j}(n_{1j}+n_{2j})}}\right)\left(t_j^2 + \dfrac{n_{2j}-n_{1j}}{n_{1j}+2n_{2j}}\right) & t_j \ge t_{\min j} \\ \\ t_j + \dfrac{g}{6}\left(\dfrac{n_{1j}+2n_{2j}}{\sqrt{n_{1j}\,n_{2j}(n_{1j}+n_{2j})}}\right)\left(t_{\min j}^2 + \dfrac{n_{2j}-n_{1j}}{n_{1j}+2n_{2j}}\right) & \text{otherwise} \end{cases},$$

where

$$t_j = \frac{\bar{X}_{1j} - \bar{X}_{2j}}{s_{1j}\sqrt{\frac{1}{n_{1j}} + \frac{1}{n_{2j}}}},$$

$$t_{\min j} = \frac{-3\sqrt{n_{1j}n_{2j}n_j}}{g(n_{1j}+2n_{2j})}$$

and g is the median value of all values of

$$\gamma_{1j} = \frac{n_{1j}}{(n_{1j}-1)(n_{1j}-2)}\sum_k\left(\frac{X_{1jk}-\bar{X}_{1j}}{s_{1j}}\right)^3$$

with $n_{1j} > n_{3q}$ for all values of *j*. n_{3q} is the 3 quartile of all values of n_{1j}.

Note, that t_j is the "modified Z" statistic. The statistic T_j is a "modified Z" corrected for the skewness of the ILEC data.

If $\min(n_{1j}, n_{2j}) \le 6$, and

a) $M_j \le 1{,}000$ (the total number of distinct pairs of samples of size n_{1j} and n_{2j} is 1,000 or less).

- Calculate the sample sum for all possible samples of size n_{2j}.
- Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
- Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\alpha = 1 - \frac{R_0 - 0.5}{M_j}$$

b) $M_j > 1{,}000$

- Draw a random sample of 1,000 sample sums from the permutation distribution.
- Add the observed sample sum to the list. There are a total of 1001 sample sums. Rank the sample sums from smallest to largest. Ties are dealt by using average ranks.
- Let R_0 be the rank of the observed sample sum with respect all the sample sums.

$$\alpha = 1 - \frac{R_0 - 0.5}{1001}.$$

Proportion Measure

$$Z_j = \frac{n_j\, a_{1j} - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\,(n_j - a_j)}{n_j - 1}}}.$$

Rate Measure

$$Z_j = \frac{n_{1j} - n_j\, q_j}{\sqrt{n_j\, q_j (1 - q_j)}}$$

Ratio Measure

$$Z_j = \frac{\hat{R}_{1j} - \hat{R}_{2j}}{\sqrt{V(\hat{R}_{1j})\left(\frac{1}{n_{1j}} + \frac{1}{n_{2j}}\right)}}$$

$$V(\hat{R}_{1j}) = \frac{\sum_k \left(U_{1jk} - \hat{R}_{1j}X_{1jk}\right)^2}{\bar{X}_{1j}^2(n_{1j}-1)} = \frac{\sum_k U_{1jk}^2 - 2\hat{R}_{1j}\sum_k \left(U_{1jk}X_{1jk}\right) + \hat{R}_{1j}^2 \sum_k X_{1jk}^2}{\bar{X}_{1j}^2(n_{1j}-1)}$$

3. **Obtain a truncated Z value for each cell,** Z_j^*. To limit the amount of cancellation that takes place between cell results during aggregation, cells whose results suggest possible favoritism are left alone. Otherwise the cell statistic is set to zero. This means that positive equivalent Z values are set to 0, and negative values are left alone. Mathematically, this is written as

$$Z_j^* = \min(0, Z_j).$$

4. **Calculate the theoretical mean and variance of the truncated statistic under the null hypothesis of parity,** $E(Z_j^*|H_0)$ and $Var(Z_j^*|H_0)$. In order to compensate for the truncation in step 3, an aggregated, weighted sum of the Z_j^* will need to be centered and scaled properly so that the final aggregate statistic follows a standard normal distribution.

   - If $W_j = 0$, then no evidence of favoritism is contained in the cell. The formulae for calculating $E(Z_j^* \mid H_0)$ and $Var(Z_j^* \mid H_0)$ cannot be used. Set both equal to 0.

   - If $\min(n_{1j}, n_{2j}) > 6$ for a mean measure, $\min\left\{a_{1j}\left(1-\frac{a_{1j}}{n_{1j}}\right), a_{2j}\left(1-\frac{a_{2j}}{n_{2j}}\right)\right\} > 9$ for a proportion measure, $\min\left(n_{1j}, n_{2j}\right) > 15$ and $n_j q_j (1-q_j) > 9$ for a rate measure, or n_{1j} and n_{2j} are large for a ratio measure then

$$E(Z_j^* \mid H_0) = -\frac{1}{\sqrt{2\pi}}, \text{ and}$$

$$Var(Z_j^* \mid H_0) = \frac{1}{2} - \frac{1}{2\pi}.$$

- Otherwise, determine the total number of values for Z_j^*. Let z_{ji} and θ_{ji}, denote the values of Z_j^* and the probabilities of observing each value, respectively.

$$E(Z_j^* \mid H_0) = \sum_i \theta_{ji} z_{ji} \text{ ,and}$$
$$Var(Z_j^* \mid H_0) = \sum_i \theta_{ji} z_{ji}^2 - \left[E(Z_j^* \mid H_0)\right]^2 .$$

The actual values of the z's and θ's depends on the type of measure.

Mean Measure

$$N_j = \min(M_j, 1{,}000), \ \ i = 1, \ldots, N_j$$
$$z_{ji} = \min\left\{0, \Phi^{-1}\left(1 - \tfrac{R_i - 0.5}{N_j}\right)\right\} \quad \text{where } R_i \text{ is the rank of sample sum i}$$
$$\theta_j = \frac{1}{N_j}$$

Proportion Measure

$$z_{ji} = \min\left\{0, \frac{n_j\, i - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)}{n_j - 1}}}\right\}, \quad i = \max(0, a_j - n_{2j}), \ldots, \min(a_j, n_{1j})$$
$$\theta_{ji} = HG(i)$$

Rate Measure

$$z_{ji} = \min\left\{0, \frac{i - n_j\, q_j}{\sqrt{n_j\, q_j (1 - q_j)}}\right\}, \quad i = 0, \ldots, n_j$$
$$\theta_{ji} = BN(i)$$

Ratio Measure

The performance measure that is in this class is billing accuracy. If a parity test were used, the sample sizes for this measure are quite large, so there is no need for a small sample technique. If one does need a small sample technique, then a re-sampling method can be used.

1. **Calculate the aggregate test statistic, Z^T.**

$$Z^T = \frac{\sum_j W_j Z_j^* - \sum_j W_j E(Z_j^* \mid H_0)}{\sqrt{\sum_j W_j^2 \mathrm{Var}(Z_j^* \mid H_0)}}$$

The Balancing Critical Value

There are four key elements of the statistical testing process:

1. the null hypothesis, H_0, that parity exists between ILEC and CLEC services
2. the alternative hypothesis, H_a, that the ILEC is giving better service to its own customers
3. the Truncated Z test statistic, Z^T, and
4. a critical value, c

The decision rule[2] is

- If $Z^T < c$ then accept H_a.
- If $Z^T \geq c$ then accept H_0.

There are two types of error possible when using such a decision rule:

Type I Error: Deciding favoritism exists when there is, in fact, no favoritism.
Type II Error: Deciding parity exists when there is, in fact, favoritism.

The probabilities of each type of each are:

Type I Error: $\alpha = P(Z^T < c \mid H_0)$.
Type II Error: $\beta = P(Z^T \geq c \mid H_a)$.

We want a balancing critical value, c_B, so that $\alpha = \beta$.

It can be shown that.

[2] This decision rule assumes that a negative test statistic indicates poor service for the CLEC customer. If the opposite is true, then reverse the decision rule.

$$c_B = \frac{\sum_j W_j M(m_j, se_j) - \sum_j W_j \frac{-1}{\sqrt{2\pi}}}{\sqrt{\sum_j W_j^2 V(m_j, se_j)} + \sqrt{\sum_j W_j^2 \left(\frac{1}{2} - \frac{1}{2\pi}\right)}}.$$

where

$$M(\mu,\sigma) = \mu\,\Phi(\tfrac{-\mu}{\sigma}) - \sigma\,\phi(\tfrac{-\mu}{\sigma})$$

$$V(\mu,\sigma) = (\mu^2 + \sigma^2)\Phi(\tfrac{-\mu}{\sigma}) - \mu\,\sigma\,\phi(\tfrac{-\mu}{\sigma}) - M(\mu,\sigma)^2$$

$\Phi(\cdot)$ is the cumulative standard normal distribution function, and $\phi(\cdot)$ is the standard normal density function.

This formula assumes that Z_j is approximately normally distributed within cell j. When the cell sample sizes, n_{1j} and n_{2j}, are small this may not be true. It is possible to determine the cell mean and variance under the null hypothesis when the cell sample sizes are small. It is much more difficult to determine these values under the alternative hypothesis. Since the cell weight, W_j will also be small (see calculate weights section above) for a cell with small volume, the cell mean and variance will not contribute much to the weighted sum. Therefore, the above formula provides a reasonable approximation to the balancing critical value.

The values of m_j and se_j will depend on the type of performance measure.

Mean Measure

For mean measures, one is concerned with two parameters in each cell, namely, the mean and variance. A possible lack of parity may be due to a difference in cell means, and/or a difference in cell variances. One possible set of hypotheses that capture this notion, and take into account the assumption that transaction are identically distributed within cells is:

$$H_0: \mu_{1j} = \mu_{2j},\ \sigma_{1j}^2 = \sigma_{2j}^2$$

$$H_a: \mu_{2j} = \mu_{1j} + \delta_j \cdot \sigma_{1j},\ \sigma_{2j}^2 = \lambda_j \cdot \sigma_{1j}^2 \qquad \delta_j > 0,\ \lambda_j \geq 1 \text{ and } j = 1,\ldots,L.$$

Under this form of alternative hypothesis, the cell test statistic Z_j has mean and standard error given by

$$m_j = \frac{-\delta_j}{\sqrt{\frac{1}{n_{1j}} + \frac{1}{n_{2j}}}}, \text{ and}$$

$$se_j = \sqrt{\frac{\lambda_j n_{1j} + n_{2j}}{n_{1j} + n_{2j}}}$$

Proportion Measure

For a proportion measure there is only one parameter of interest in each cell, the proportion of transaction possessing an attribute of interest. A possible lack of parity may be due to a difference in cell proportions. A set of hypotheses that take into account the assumption that transaction are identically distributed within cells while allowing for an analytically tractable solution is:

$$H_0: \frac{p_{2j}(1-p_{1j})}{(1-p_{2j})p_{1j}} = 1$$

$$H_a: \frac{p_{2j}(1-p_{1j})}{(1-p_{2j})p_{1j}} = \psi_j \qquad \psi_j > 1 \text{ and } j = 1,\ldots,L.$$

These hypotheses are based on the "odds ratio." If the transaction attribute of interest is a missed trouble repair, then an interpretation of the alternative hypothesis is that a CLEC trouble repair appointment is ψ_j times more likely to be missed than an ILEC trouble.

Under this form of alternative hypothesis, the within cell asymptotic mean and variance of a_{1j} are given by[3]

$$E(a_{1j}) = n_j \pi_j^{(1)}$$

$$\text{var}(a_{1j}) = \frac{n_j}{\frac{1}{\pi_j^{(1)}} + \frac{1}{\pi_j^{(2)}} + \frac{1}{\pi_j^{(3)}} + \frac{1}{\pi_j^{(4)}}}$$

where

[3] Stevens, W. L. (1951) Mean and Variance of an entry in a Contingency Table. *Biometrica*, **38**, 468-470.

$$\pi_j^{(1)} = f_j^{(1)}\left(n_j^2 + f_j^{(2)} + f_j^{(3)} - f_j^{(4)}\right)$$
$$\pi_j^{(2)} = f_j^{(1)}\left(-n_j^2 - f_j^{(2)} + f_j^{(3)} + f_j^{(4)}\right)$$
$$\pi_j^{(3)} = f_j^{(1)}\left(-n_j^2 + f_j^{(2)} - f_j^{(3)} + f_j^{(4)}\right)$$
$$\pi_j^{(4)} = f_j^{(1)}\left(n_j^2\left(\tfrac{2}{\psi_j} - 1\right) - f_j^{(2)} - f_j^{(3)} - f_j^{(4)}\right)$$
$$f_j^{(1)} = \frac{1}{2n_j^2\left(\frac{1}{\psi_j} - 1\right)}$$
$$f_j^{(2)} = n_j n_{1j}\left(\tfrac{1}{\psi_j} - 1\right)$$
$$f_j^{(3)} = n_j a_j\left(\tfrac{1}{\psi_j} - 1\right)$$
$$f_j^{(4)} = \sqrt{n_j^2\left[4n_{1j}\left(n_j - a_j\right)\left(\tfrac{1}{\psi_j} - 1\right) + \left(n_j + \left(a_j - n_{1j}\right)\left(\tfrac{1}{\psi_j} - 1\right)\right)^2\right]}$$

Recall that the cell test statistic is given by

$$Z_j = \frac{n_j\, a_{1j} - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)}{n_j - 1}}}.$$

Using the equations above, we see that Z_j has mean and standard error given by

$$m_j = \frac{n_j^2 \pi_j^{(1)} - n_{1j}\, a_j}{\sqrt{\dfrac{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)}{n_j - 1}}}, \text{ and}$$

$$se_j = \sqrt{\frac{n_j^3 (n_j - 1)}{n_{1j}\, n_{2j}\, a_j\, (n_j - a_j)\left(\frac{1}{\pi_j^{(1)}} + \frac{1}{\pi_j^{(2)}} + \frac{1}{\pi_j^{(3)}} + \frac{1}{\pi_j^{(4)}}\right)}}.$$

Rate Measure

A rate measure also has only one parameter of interest in each cell, the rate at which a phenomenon is observed relative to a base unit, e.g. the number of troubles per available line. A possible lack of parity may be due to a difference in cell rates. A set of hypotheses that take into account the assumption that transaction are identically distributed within cells is:

H_0: $r_{1j} = r_{2j}$

H_a: $r_{2j} = \varepsilon_j r_{1j}$ $\quad\quad$ $\varepsilon_j > 1$ and $j = 1,\ldots,L$.

Given the total number of ILEC and CLEC transactions in a cell, n_j, and the number of base elements, b_{1j} and b_{2j}, the number of ILEC transaction, n_{1j}, has a binomial distribution from n_j trials and a probability of

$$q_j^* = \frac{r_{1j}b_{1j}}{r_{1j}b_{1j} + r_{2j}b_{2j}}.$$

Therefore, the mean and variance of n_{1j}, are given by

$$E(n_{1j}) = n_j q_j^*$$
$$\mathrm{var}(n_{1j}) = n_j q_j^*(1 - q_j^*)$$

Under the null hypothesis

$$q_j^* = q_j = \frac{b_{1j}}{b_j},$$

but under the alternative hypothesis

$$q_j^* = q_j^a = \frac{b_{1j}}{b_{1j} + \varepsilon_j b_{2j}}.$$

Recall that the cell test statistic is given by

$$Z_j = \frac{n_{1j} - n_j\, q_j}{\sqrt{n_j\, q_j(1 - q_j)}}.$$

Using the relationships above, we see that Z_j has mean and standard error given by

$$m_j = \frac{n_j\left(q_j^a - q_j\right)}{\sqrt{n_j q_j(1 - q_j)}} = (1 - \varepsilon_j)\frac{\sqrt{n_j b_{1j} b_{2j}}}{b_{1j} + \varepsilon_j b_{2j}}, \text{ and}$$

$$se_j = \sqrt{\frac{q_j^a(1 - q_j^a)}{q_j(1 - q_j)}} = \sqrt{\varepsilon_j}\,\frac{b_j}{b_{1j} + \varepsilon_j b_{2j}}.$$

Ratio Measure

As with mean measures, one is concerned with two parameters in each cell, the mean and variance, when testing for parity of ratio measures. As long as sample sizes are large, as in the case of billing accuracy, the same method for finding m_j and se_j that is used for mean measures can be used for ratio measures.

Determining the Parameters of the Alternative Hypothesis

In this exhibit we have indexed the alternative hypothesis of mean measures by two sets of parameters, λ_j and δ_j. Proportion and rate measures have been indexed by one set of parameters each, ψ_j and ε_j respectively. A major difficulty with this approach is that more than one alternative will be of interest; for example we may consider one alternative in which all the δ_j are set to a common non-zero value, and another set of alternatives in each of which just one δ_j is non-zero, while all the rest are zero. There are very many other possibilities. Each possibility leads to a single value for the balancing critical value; and each possible critical value corresponds to many sets of alternative hypotheses, for each of which it constitutes the correct balancing value.

The formulas we have presented can be used to evaluate the impact of different choices of the overall critical value. For each putative choice, we can evaluate the set of alternatives for which this is the correct balancing value. While statistical science can be used to evaluate the impact of different choices of these parameters, there is not much that an appeal to statistical principles can offer in directing specific choices. Specific choices are best left to telephony experts. Still, it is possible to comment on some aspects of these choices:

- Parameter Choices for λ_j. The set of parameters λ_j index alternatives to the null hypothesis that arise because there might be greater unpredictability or variability in the delivery of service to a CLEC customer over that which would be achieved for an otherwise comparable ILEC customer. While concerns about differences in the variability of service are important, it turns out that the truncated Z testing which is being recommended here is relatively insensitive to all but very large values of the λ_j. Put another way, reasonable differences in the values chosen here could make very little difference in the balancing points chosen.

- Parameter Choices for δ_j. The set of parameters δ_j are much more important in the choice of the balancing point than was true for the λ_j. The reason for this is that they directly index differences in average service. The truncated Z test is very sensitive to any such differences; hence, even small disagreements among experts in the choice of the δ_j could be very important. Sample size matters here too. For example, setting all the δ_j to a single value – $\delta_j = \delta$ – might be fine for tests across individual CLECs where currently in Louisiana the CLEC customer bases are not too different. Using the same value of δ for the overall state testing does not seem sensible. At the state level we are aggregating over CLECs, so using the same δ as for an individual CLEC would be saying that a "meaningful" degree of disparity is one where the violation is the same (δ) for each CLEC. But the detection of disparity for any component CLEC is important, so the relevant "overall" δ should be smaller.

- Parameter Choices for ψ_j or ε_j. The set of parameters ψ_j or ε_j are also important in the choice of the balancing point for tests of their respective measures. The reason for this is that they directly index increases in the proportion or rate of service performance. The truncated Z test is sensitive to such increases; but not as sensitive as the case of δ for mean measures. Sample size matters here too. As with mean measures, using the same value of ψ or ε for the overall state testing does not seem sensible.

The three parameters are related however. If a decision is made on the value of δ, it is possible to determine equivalent values of ψ and ε. The following equations, in conjunction with the definitions of ψ and ε, show the relationship with delta.

$$\delta = 2 \cdot \arcsin(\sqrt{\hat{p}_2}) - 2 \cdot \arcsin(\sqrt{\hat{p}_1})$$
$$\delta = 2\sqrt{\hat{r}_2} - 2\sqrt{\hat{r}_1}$$

The bottom line here is that beyond a few general considerations, like those given above, a principled approach to the choice of the alternative hypotheses to guard against must come from elsewhere.

Decision Process

Once Z^T has been calculated, it is compared to the balancing critical value to determine if the ILEC is favoring its own customers over a CLEC's customers.

This critical value changes as the ILEC and CLEC transaction volume change. One way to make this transparent to the decision-maker, is to report the difference between the test statistic and the critical value, *diff* = Z^T - c_B. If favoritism is concluded when $Z^T < c_B$, then the *diff* < 0 indicates favoritism.

This makes it very easy to determine favoritism: a positive *diff* suggests no favoritism, and a negative *diff* suggests favoritism.